

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/S12,570
Source: PCT
Date Processed by STIC: 12-20-04

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/517,510

TIME: 12:30:52

Input Set : A:\21023P SEQUENCE 120204.TXT

Output Set: N:\CRF4\12202004\J517510.raw

```

4 <110> APPLICANT: Harvey, Diane Marie
5      Yang, Yi
6      Kohl, Nancy
8 <120> TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING
9      A NOVEL CENTROMERE-ASSOCIATED MOTOR PROTEIN, AND USES
10     THEREOF
12 <130> FILE REFERENCE: 21023P
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/517,510
C--> 14 <141> CURRENT FILING DATE: 2004-12-10
14 <150> PRIOR APPLICATION NUMBER: PCT/US03/18203
15 <151> PRIOR FILING DATE: 2003-06-09
17 <150> PRIOR APPLICATION NUMBER: 60/387,403
18 <151> PRIOR FILING DATE: 2002-06-10
20 <160> NUMBER OF SEQ ID NOS: 4
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 1488
26 <212> TYPE: DNA
27 <213> ORGANISM: Human
29 <400> SEQUENCE: 1
30 atggcggagg aaggagccgt ggccgtctgc gtgcgagtg ggcgcgtgaa cagcagagaa 60
31 gaatcacttg gagaaactgc ccaagtttac tggaaaaactg catttatcaa 120
32 gttgatggaa gtaaatcctt caattttgat cgtgtctttc atggtaatga aactaccaa 180
33 aatgtgtatg aagaaatagc agcaccaatc atcgattctg ccatacaagg ctacaatggt 240
34 actatatttg cctatggaca gactgcttca ggaaaaacat ataccatgat gggttcagaa 300
35 gatcatttgg gagttatacc cagggcaatt catgacattt tccaaaaaat taagaagttt 360
36 cctgataggg aatttctctt acgtgtatct tacatggaaa tatacaatga aaccattaca 420
37 gattttactct gtggcactca aaaaatgaaa cctttaatta ttcgagaaga tgtcaatagg 480
38 aatgtgtatg ttgctgatct cacagaagaa gttgtatata catcagaaat ggctttgaaa 540
39 tggattacaa agggagaaaa gagcaggcat tatggagaaa caaaaatgaa tcaaagaagc 600
40 agtcgttctc ataccatctt taggatgatt ttggaaagca gagagaaggg tgaaccttct 660
41 aattgtgaag gatctgttaa ggtatcccat ttgaatttgg ttgatcttgc aggcagtga 720
42 agagctgctc aaacaggcgc tgcagggtgtg cggtcaagg aaggctgtaa tataaatcga 780
43 agcttattta ttttgggaca agtgatcaag aaacttagtg atggacaagt tggtggtttc 840
44 ataaattatc gagatagcaa gttaacacga attctccaga attccttggg aggaaatgca 900
45 aagacacgta ttatctgcac aattactcca gtatcttttg atgaaacact tactgctctc 960
46 cagtttgcca gtactgctaa atatatgaag aatactcctt atgttaatga ggtatcaact 1020
47 gatgaagctc tcctgaaaag gtatagaaaa gaaataatgg atcttaaaaa acaattagag 1080
48 gaggtttctt tagagacgcy ggctcaggca atggaaaaag accaattggc ccaacttttg 1140
49 gaagaaaaag atttgcttca gaaagtacag aatgagaaaa ttgaaaactt aacacggaag 1200
50 ctgggtgacct cttcttccct caggttgcaa caggaattaa aggctaaaag aaaacgaaga 1260
51 gttacttggt gccttgcaa aattaacaaa atgaagaact caaactatgc agatcaattt 1320
52 aatataccaa caaatataac aacaaaaaca cataagcttt ctataaattt attacgagaa 1380

```

RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/517,510

TIME: 12:30:52

Input Set : A:\21023P SEQUENCE 120204.TXT

Output Set: N:\CRF4\12202004\J517510.raw

```

53 attgatgaat ctgtcgggtac catatgggaa ttcgaagctt acgtagaaca aaaactcatc 1440
54 tcagaagagg atctgaatag cgccgtcgac catcatcatc atcatcat 1488
56 <210> SEQ ID NO: 2
57 <211> LENGTH: 496
58 <212> TYPE: PRT
59 <213> ORGANISM: Human
61 <400> SEQUENCE: 2
62 Met Ala Glu Glu Gly Ala Val Ala Val Cys Val Arg Val Arg Pro Leu
63 1 5 10 15
64 Asn Ser Arg Glu Ser Leu Gly Glu Thr Ala Gln Val Tyr Trp Lys
65 20 25 30
66 Thr Asp Asn Asn Val Ile Tyr Gln Val Asp Gly Ser Lys Ser Phe Asn
67 35 40 45
68 Phe Asp Arg Val Phe His Gly Asn Glu Thr Thr Lys Asn Val Tyr Glu
69 50 55 60
70 Glu Ile Ala Ala Pro Ile Ile Asp Ser Ala Ile Gln Gly Tyr Asn Gly
71 65 70 75 80
72 Thr Ile Phe Ala Tyr Gly Gln Thr Ala Ser Gly Lys Thr Tyr Thr Met
73 85 90 95
74 Met Gly Ser Glu Asp His Leu Gly Val Ile Pro Arg Ala Ile His Asp
75 100 105 110
76 Ile Phe Gln Lys Ile Lys Lys Phe Pro Asp Arg Glu Phe Leu Leu Arg
77 115 120 125
78 Val Ser Tyr Met Glu Ile Tyr Asn Glu Thr Ile Thr Asp Leu Leu Cys
79 130 135 140
80 Gly Thr Gln Lys Met Lys Pro Leu Ile Ile Arg Glu Asp Val Asn Arg
81 145 150 155 160
82 Asn Val Tyr Val Ala Asp Leu Thr Glu Glu Val Val Tyr Thr Ser Glu
83 165 170 175
84 Met Ala Leu Lys Trp Ile Thr Lys Gly Glu Lys Ser Arg His Tyr Gly
85 180 185 190
86 Glu Thr Lys Met Asn Gln Arg Ser Ser Arg Ser His Thr Ile Phe Arg
87 195 200 205
88 Met Ile Leu Glu Ser Arg Glu Lys Gly Glu Pro Ser Asn Cys Glu Gly
89 210 215 220
90 Ser Val Lys Val Ser His Leu Asn Leu Val Asp Leu Ala Gly Ser Glu
91 225 230 235 240
92 Arg Ala Ala Gln Thr Gly Ala Ala Gly Val Arg Leu Lys Glu Gly Cys
93 245 250 255
94 Asn Ile Asn Arg Ser Leu Phe Ile Leu Gly Gln Val Ile Lys Lys Leu
95 260 265 270
96 Ser Asp Gly Gln Val Gly Gly Phe Ile Asn Tyr Arg Asp Ser Lys Leu
97 275 280 285
98 Thr Arg Ile Leu Gln Asn Ser Leu Gly Gly Asn Ala Lys Thr Arg Ile
99 290 295 300
100 Ile Cys Thr Ile Thr Pro Val Ser Phe Asp Glu Thr Leu Thr Ala Leu
101 305 310 315 320
102 Gln Phe Ala Ser Thr Ala Lys Tyr Met Lys Asn Thr Pro Tyr Val Asn
103 325 330 335

```

RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/517,510

TIME: 12:30:52

Input Set : A:\21023P SEQUENCE 120204.TXT

Output Set: N:\CRF4\12202004\J517510.raw

```

104 Glu Val Ser Thr Asp Glu Ala Leu Leu Lys Arg Tyr Arg Lys Glu Ile
105           340           345           350
106 Met Asp Leu Lys Lys Gln Leu Glu Glu Val Ser Leu Glu Thr Arg Ala
107           355           360           365
108 Gln Ala Met Glu Lys Asp Gln Leu Ala Gln Leu Leu Glu Glu Lys Asp
109           370           375           380
110 Leu Leu Gln Lys Val Gln Asn Glu Lys Ile Glu Asn Leu Thr Arg Met
111 385           390           395           400
112 Leu Val Thr Ser Ser Ser Leu Thr Leu Gln Gln Glu Leu Lys Ala Lys
113           405           410           415
114 Arg Lys Arg Arg Val Thr Trp Cys Leu Gly Lys Ile Asn Lys Met Lys
115           420           425           430
116 Asn Ser Asn Tyr Ala Asp Gln Phe Asn Ile Pro Thr Asn Ile Thr Thr
117           435           440           445
118 Lys Thr His Lys Leu Ser Ile Asn Leu Leu Arg Glu Ile Asp Glu Ser
119           450           455           460
120 Val Gly Thr Ile Trp Glu Phe Glu Ala Tyr Val Glu Gln Lys Leu Ile
121 465           470           475           480
122 Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His His His His His
123           485           490           495
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 1065
128 <212> TYPE: DNA
129 <213> ORGANISM: Human
131 <400> SEQUENCE: 3
132 atggcgaggagg aaggagccgt ggccgtctgc gtgcgagtgcc ggccgctgaa cagcagagaa 60
133 gaatcacttg gagaaactgc ccaagtttac tggaaaactg acaataatgt catttatcaa 120
134 gttgatggaa gtaaatcctt caattttgat cgtgtctttc atggtaatga aactaccaa 180
135 aatgtgtatg aagaaatagc agcaccaatc atcgattctg ccatacaagg ctacaatgg 240
136 actatatttg cctatggaca gactgcttca ggaaaaacat ataccatgat ggggttcagaa 300
137 gatcatttgg gagttatacc cagggcaatt catgacattt tccaaaaaat taagaagttt 360
138 cctgataggg aatttctctt acgtgtatct tacatggaaa tatacaatga aaccattaca 420
139 gatttactct gtggcactca aaaaatgaaa cctttaatta ttcgagaaga tgtcaatagg 480
140 aatgtgtatg ttgctgatct cacagaagaa gttgtatata catcagaaat ggctttgaaa 540
141 tggattacaa agggagaaaa gagcaggcat tatggagaaa caaaaatgaa tcaaagaagc 600
142 agtcgttctc ataccatctt taggatgatt ttggaaagca gagagaaggg tgaaccttct 660
143 aattgtgaag gatctgttaa ggtatcccat ttgaatttgg ttgatcttgc aggcagttaa 720
144 agagctgctc aaacaggcgc tgcagggtgt cggtcgaagg aaggctgtaa tataaatcga 780
145 agcttattta ttttgggaca agtgatcaag aaacttagtg atggacaagt tgggtggtttc 840
146 ataaattatc gagatagcaa gttaacacga attctccaga attccttggg aggaaatgca 900
147 aagacacgta ttatctgcac aattactcca gtatcttttg atgaaacact tactgctctc 960
148 cagtttgcca gtactgctaa atatatgaag aatactcctt atgttaatga ggtatcaact 1020
149 gtcgacaagc ttgcggccgc actcgagcac caccaccacc accac 1065
151 <210> SEQ ID NO: 4
152 <211> LENGTH: 355
153 <212> TYPE: PRT
154 <213> ORGANISM: Human
156 <400> SEQUENCE: 4
157 Met Ala Glu Glu Gly Ala Val Ala Val Cys Val Arg Val Arg Pro Leu

```

RAW SEQUENCE LISTING

DATE: 12/20/2004

PATENT APPLICATION: US/10/517,510

TIME: 12:30:52

Input Set : A:\21023P SEQUENCE 120204.TXT

Output Set: N:\CRF4\12202004\J517510.raw

```

158 1          5          10          15
159 Asn Ser Arg Glu Glu Ser Leu Gly Glu Thr Ala Gln Val Tyr Trp Lys
160          20          25          30
161 Thr Asp Asn Asn Val Ile Tyr Gln Val Asp Gly Ser Lys Ser Phe Asn
162          35          40          45
163 Phe Asp Arg Val Phe His Gly Asn Glu Thr Thr Lys Asn Val Tyr Glu
164          50          55          60
165 Glu Ile Ala Ala Pro Ile Ile Asp Ser Ala Ile Gln Gly Tyr Asn Gly
166 65          70          75          80
167 Thr Ile Phe Ala Tyr Gly Gln Thr Ala Ser Gly Lys Thr Tyr Thr Met
168          85          90          95
169 Met Gly Ser Glu Asp His Leu Gly Val Ile Pro Arg Ala Ile His Asp
170          100          105          110
171 Ile Phe Gln Lys Ile Lys Lys Phe Pro Asp Arg Glu Phe Leu Leu Arg
172          115          120          125
173 Val Ser Tyr Met Glu Ile Tyr Asn Glu Thr Ile Thr Asp Leu Leu Cys
174          130          135          140
175 Gly Thr Gln Lys Met Lys Pro Leu Ile Ile Arg Glu Asp Val Asn Arg
176 145          150          155          160
177 Asn Val Tyr Val Ala Asp Leu Thr Glu Glu Val Val Tyr Thr Ser Glu
178          165          170          175
179 Met Ala Leu Lys Trp Ile Thr Lys Gly Glu Lys Ser Arg His Tyr Gly
180          180          185          190
181 Glu Thr Lys Met Asn Gln Arg Ser Ser Arg Ser His Thr Ile Phe Arg
182          195          200          205
183 Met Ile Leu Glu Ser Arg Glu Lys Gly Glu Pro Ser Asn Cys Glu Gly
184          210          215          220
185 Ser Val Lys Val Ser His Leu Asn Leu Val Asp Leu Ala Gly Ser Glu
186 225          230          235          240
187 Arg Ala Ala Gln Thr Gly Ala Ala Gly Val Arg Leu Lys Glu Gly Cys
188          245          250          255
189 Asn Ile Asn Arg Ser Leu Phe Ile Leu Gly Gln Val Ile Lys Lys Leu
190          260          265          270
191 Ser Asp Gly Gln Val Gly Gly Phe Ile Asn Tyr Arg Asp Ser Lys Leu
192          275          280          285
193 Thr Arg Ile Leu Gln Asn Ser Leu Gly Gly Asn Ala Lys Thr Arg Ile
194          290          295          300
195 Ile Cys Thr Ile Thr Pro Val Ser Phe Asp Glu Thr Leu Thr Ala Leu
196 305          310          315          320
197 Gln Phe Ala Ser Thr Ala Lys Tyr Met Lys Asn Thr Pro Tyr Val Asn
198          325          330          335
199 Glu Val Ser Thr Val Asp Lys Leu Ala Ala Ala Leu Glu His His His
200          340          345          350
201 His His His
202          355

```

VERIFICATION SUMMARY

DATE: 12/20/2004

PATENT APPLICATION: US/10/517,510

TIME: 12:30:53

Input Set : A:\21023P SEQUENCE 120204.TXT

Output Set: N:\CRF4\12202004\J517510.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date